

***Would have gotten away
with it, if it hadn't been for
you meddling librarians:
Unmasking AI ghost and
phantom references***

Dr Fred Milton

Liaison Librarian, Teesside University Library

Email: f.milton@tees.ac.uk



Introduction

Evolution of AI Large Language Models

- ChatGPT3.5 introduced Nov 2022, Scholar GPT Jan 2024 & other interactions – now ChatGPT5.5
- MS flagship AI chatbot – Bing Chat Feb 2023, rebranded as MS Copilot Nov 2023. Approved use at Teesside
- Other AI models available.

Currently at Teesside University:

- AI use by students is steered by a Red, Amber, Green traffic light system – currently under review.

New problem

- Production of confabulated, 'ghost' or 'hallucinated' citations in response to literature searching prompts
- 35% of students using generative AI unaware of how often it produces made up information (Freeman, 2024).



Definitions

Definitions

- '*hallucination*, the generated content that is nonsensical or unfaithful to the provided source content' (Ji *et al.* 2023)
- '*Ghost publication*...defined as any publication listed as 'published', 'accepted', or 'in-press' that did not exist in the literature' (ElHawary *et al.* 2022).

Growing body of research exploring proclivity of these systems to produce error-strewn and often fabricated references with heavy focus on critique of ChatGPT.

What can we do to guide AI use for literature searching?

What exactly are these AI prompts returning in terms of quality and reliability?



ChatGPT (2026) AI-generated image by ChatGPT with prompt 'cartoon image of a hallucinations', 18 June. Available at: https://chatgpt.com/s/m_6a3412c3cd34819181c118852165402c (Accessed: 18 June 2026).

A concerted search for ghosts...

Exemplar essay title used to coach Teesside University business undergraduates' literature search skills and identify key search words

The effects of viral marketing on teenagers' buying habits

AI prompt

I am doing a literature review for the follow undergraduate assignment "The effects of viral marketing on teenagers' buying habits". Please provide me 20 academic journal articles that would help me. The articles should be referenced in the Harvard reference style. Thank you

Five ChatGPT and Copilot prompts

Five Caveats

- 2 prompts with no date limitations
- 2 with filters of 'last 5 years'
- Further Chat GPT prompt with a newly created account (no date filter).



Searching for ghosts - results

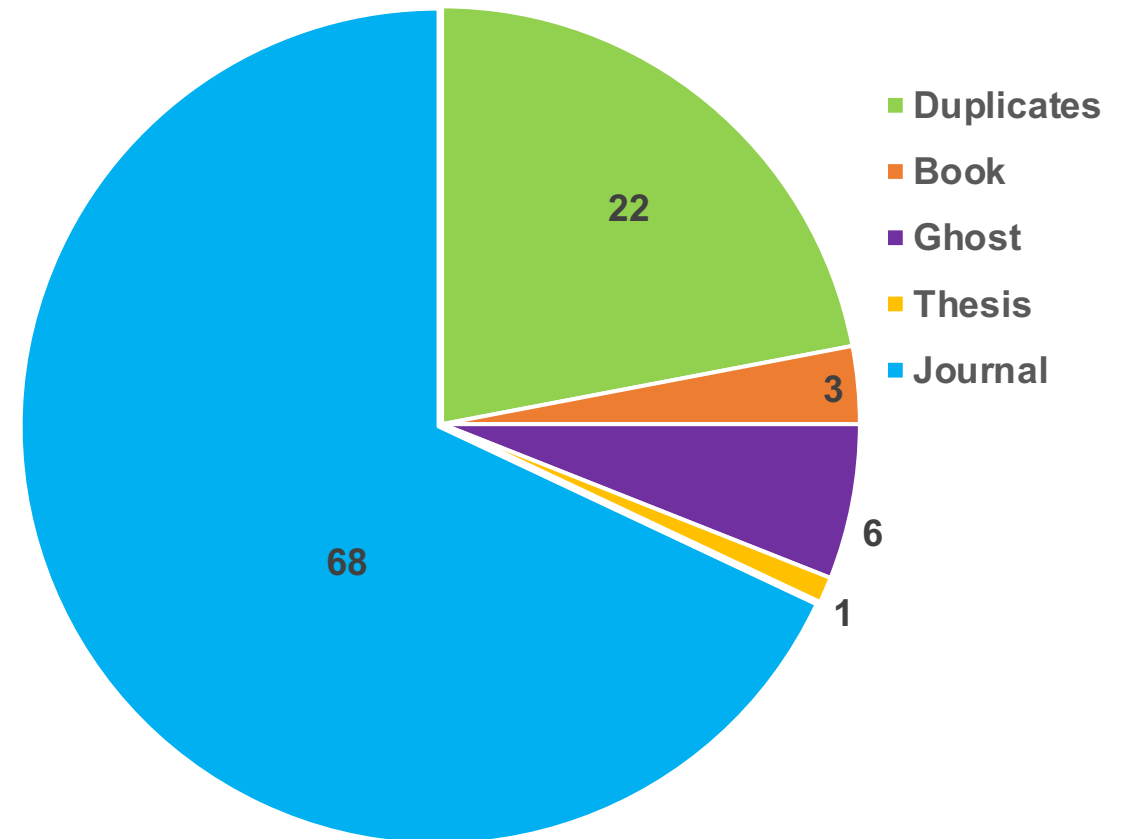
100 citations retrieved

- 22 citations duplicates - one article cited 4 times in both Copilot and Chat GPT returns
- No obvious difference between 'date' filters and 'new account' prompts.

Hallucinations/errors in Harvard referencing - incomplete citations/incorrect details

- 100 citations – 20 full and correct Harvard style
- ChatGPT provided just 6 error-free citations
- Harvard format Author, Initial. (Year of publication) 'Title of article', Title of Journal, Issue information, Page range
- DOI/URL not assessed.

Breakdown of 100 AI Citations Retrieved



Unmasking the 6 ghosts

Analysis of the 6 Ghost journal references – 3 each from Copilot & Chat GPT

- All 6 **perfectly referenced** in Harvard style
- Claim to be in **reputable journals**, but not located in these editions & page nos. don't match
- Google AI offers an **AI overview** of articles & key arguments
- **Ghosts 1 & 2** - amalgam of genuine book titles
- **Ghost 3** – amalgam of published paper
Haur, T.S., *et al.* (2017) 'Digital advertising effectiveness and youth engagement', *Journal of Marketing Communications*, 23(4), pp.1–15
Should have been Haur, F. C., Khatibi, A., & Azam, S. M. F. (2017) 'Antecedents of consumers ' perception towards online advertising in Malaysia : The structure equation modeling approach', *European Journal of Management and Marketing Studies*, 2(3), pp. 15–30.
- **Ghost 4**
Chen, X. and Zhao, Y. (2021) 'Influencer marketing and adolescent consumer behavior: evidence from Chinese social media', *Journal of Youth Studies*, 24(6), pp. 785–801
Not located in J of Y Studies. References cited in another published article & Google AI provides an AI overview.
- **Ghosts 5 & 6** - in-text citations in a published paper, but full reference not located in article reference list.

But...is there something scarier than ghosts?

Scarier than ghosts? Predatory journals

Jeffrey Beall, a scholarly communications librarian, coined the phrase 'predatory journal' in 2010 and managed a list of potential predator publishers (Chandra and Dasgupta, 2024).

Definition

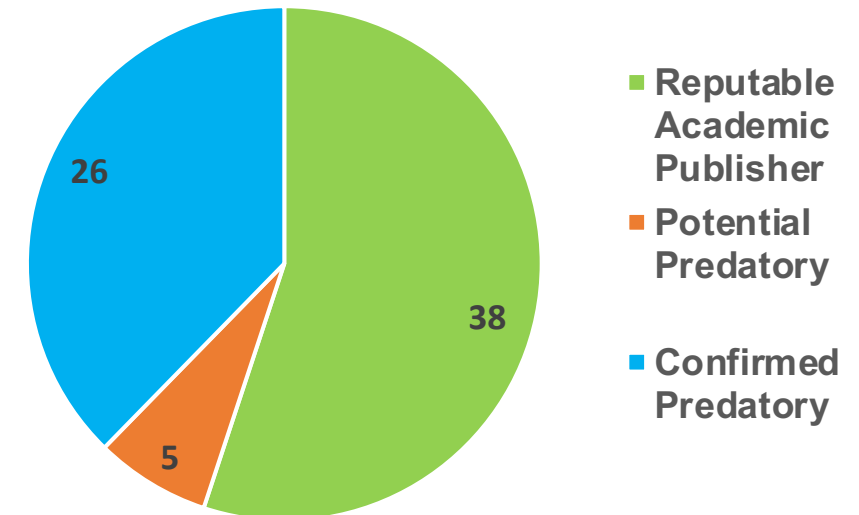
'Predatory journals and publishers are entities that prioritize self-interest at the expense of scholarship and are characterized by **false or misleading information, deviation from best editorial and publication practices**, a lack of transparency, and/or the use of **aggressive and indiscriminate practices**' (Grudniewicz *et al.*, 2019).

68 journal articles. Their host journal assessed for status via Google AI search

- **38 articles were in journals regarded as 'Reputable Publisher'** – 'a highly respected, legitimate academic publication'...institutional publisher, renowned academic publisher etc..
- **5 articles were in journals regarded as 'Potential Predatory'** - 'widely flagged by academics and librarians as a potentially predatory or low-quality journal'
- **26 were in journals regarded as 'Predatory'** - 'is widely regarded as predatory or at least seriously problematic'.

Little (if no?) academic assessment of AI searching prompts returning predatory journal articles

Analysis of 68 Journal Publication Citations



What's the problem with predators?

15,000 predatory journal titles identified in 2021 (Wilson, 2024).

- **False or misleading information** – Publishing low-quality or false research. Spoof paper Daniel Baldassarre (2020) 'What's the Deal with Birds?', *Journal of Research and Reviews*. Acknowledges 'Big Bird from Sesame Street for comments on the manuscript'.
- **Deviation from best editorial and publication practice** – Lack of retraction policy, transfer of copyright, not specifying CC licence in OA journal...
- **Lack of transparency** - Contracts, fees, charges, unverifiable editors, peer review, misleading metrics/impact factors, obfuscating titles – one journal includes a European city in its title but is published in Indonesia
- **Aggressive and indiscriminate solicitation** - Repeated soliciting and over-flattering emails.

(Grudniewicz *et al.*, 2019; Cook, Govender, and Brennan, 2023; Wilson, 2024)



Potential issues for students/researchers

Literature searching skills

- **Uncritical** reading, ease of searching/retrieval via Google and AI, time pressures, speed of returns
- **Open access** through Google/AI (many Predatory Journal articles can't be located via Google Scholar)
- **Lack of confidence** using Discovery/Databases (might improve with new natural language searching)
- MyBib suggests predatory articles are **credible**
- **Widespread** use of AI in previous institution/lack of guardrail AI guidance/policy.

Predatory journal problems – harder to identify

- Predatory journal articles **cited in published articles** given credence
- **Confusion/mimic** of established journal titles
- **Rapid or no 'peer review'** and hasty publication times suggest little or no quality peer-review (e.g. Received 12 Aug., Revised 22 Aug. & Accepted 24 Aug. 2025)
- Status and affiliation of **authors omitted/unclear**
- Massive **publication output** volumes (e.g. 1,000+ articles in one 2022 vol). Profit over academic enquiry
- Articles with **no intext citations/poor quality references**
- Some **previously legitimate journals**, available in search portals, now viewed as predatory.

Critical thinking AI retrieved sources checklist

- Does it use references? **Reputable journals, book publishers, websites...**
- Background or focussed information – **does it 'waffle'?**
- **Does the article look well formatted, 'professional' and in the journal it claims?**
- **Unrealistic broad scope of journal - accept and publish anything**
- Date of publication. Is it 'dated'?
- **Dates article received, accepted and published – feasible for a 'peer reviewed' paper and evidence of peer-review?**
- Authority of authors & affiliation – **can these be traced?**
- **Grammar and punctuation – evidence of proof reading?**
- **Arguments – does it makes sense?**
- Quantitative data - check reliability and validity
- **Source of citation? AI search, Discovery, Database, Secondary source citation... follow up any source links provided by AI results**
- Critical thinking workshop https://libguides.tees.ac.uk/critical_thinking



Applying critical thinking to ‘ghost’ articles

Produced by AI generative tool

Chen, X. and Zhao, Y. (2021) ‘Influencer marketing and adolescent consumer behavior: evidence from Chinese social media’, *Journal of Youth Studies*, 24(6), pp. 785–801

☺ Author – X. Chen – multiple scholars of same name in field of social media & business research; Yuxin Zhao - ?
Dept of Social Medicine & Health Management, Xiangya School of Public Health, Central South University, Changsha, China?

☺ Journal – Taylor & Francis peer-reviewed journal

☺ Reference – there is a 2021 vol 24/6 of journal but no paper of this name

☺ Google AI analysis & arguments

☹ Article doesn’t exist!! As Google Scholar & Discovery quickly show

- Commit to do your own searches - use our **Discovery** search portal & **databases**
- Develop your **own** original ideas & writing style – what do you think?
- See <https://libguides.tees.ac.uk/ArtificialIntelligence>

If you’ve permission to use AI, **evidence searches undertaken** to find the resources to support your assignment & bibliography. **These strategies are not about avoiding plagiarism, it’s about protecting the quality & credibility of your work.**

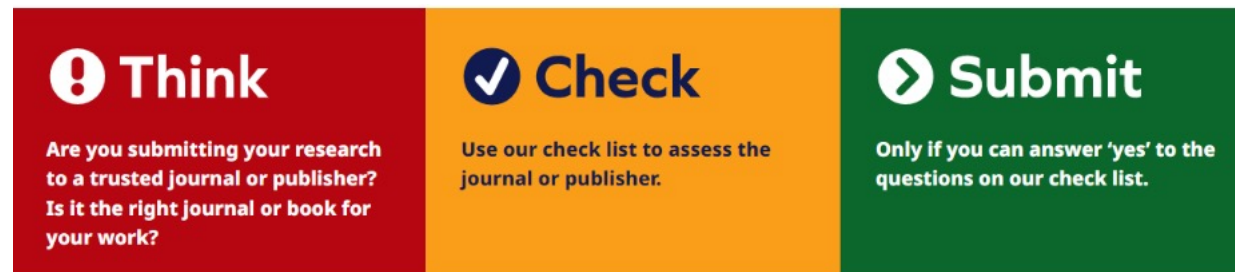
Reflection and action

Literature Searching – not all negative

- Quality of AI search returns improving following updates – 55% of GPT -3.5 citation fabricated; 18% GPT-4 fabricated (Walters and Wilder, 2023; Spennemann, 2025). **NB** – no comment about predatory journals
- Traditional web searching remains highly popular
- Students view AI as just another tool in their literature searching tool kit (Lund *et al.*, 2026).
- Reluctance of students to fully embrace AI searching - retain suspicion of AI, confidentiality, unfair and environmental objections

Raising student awareness of AI pitfalls at Teesside University

- Student participation in AI Resilient Learner Module, Ethical and Responsible AI Use training and new Critical AI Library tutorials
- Integration into Library Academic Skills workshops - searches in databases will always retrieve the exact same paper
- Existing advice on predatory journals for PG/staff potential publishers to be flagged to all students
https://libguides.tees.ac.uk/researchers/predatory_publishers



- AI Library workshop <https://libguides.tees.ac.uk/ArtificialIntelligence> and AI awareness integrated into all other Library workshops

Where next?

Where next for this research?

- More incisive detailed exemplar question for AI prompt to better critique AI responses and Test Scholar ChatGPT
- Critical read of AI search-returned articles – errors, typos, arguments, grammar, sources, data...
- Evaluation of student submitted reference/bibliographies – requires TU ethics agreement
- Potential journal article – heavy academic focus in sciences and health subjects, not so business
- Peer feedback – BLA and ALISS conferences



References

- Cook, F., Govender, R. and Brennan, P.A. (2023) 'Greetings from your predatory journal! What they are, why they are a problem, how to spot and avoid them,' *British journal of oral and maxillofacial surgery*, 61(3), pp. 245–247. Available at: <https://doi.org/10.1016/j.bjoms.2023.02.005>.
- Chandra, A. and Dasgupta, S. (2024) 'Predatory Journals: What the Researchers and Authors Should Know,' *The American journal of medicine*, 137(6), pp. 470–472. Available at: <https://doi.org/10.1016/j.amjmed.2024.02.015>.
- Freeman, J. (2024) *Provide or punish? Students' views on generative AI in higher education*. Available at: <https://www.hepi.ac.uk/wp-content/uploads/2024/01/HEPI-Policy-Note-51.pdf>. (Accessed: 5 June 2026).
- Grudniewicz, A. *et al.* (2019) 'Predatory journals: no definition, no defence,' *Nature*, 576(7786), pp. 210–212. Available at: <https://doi.org/10.1038/d41586-019-03759-y>.
- Lund, B.D *et al.* (2026) 'Artificial intelligence (AI) and information seeking: A comparative exploration of AI chatbots, search engines, and library resources as information sources among university students,' *Journal of librarianship and information science* [Preprint]. Available at: <https://doi.org/10.1177/09610006261438484>.
- Oermann, M.H. *et al.* (2019) 'Citations of articles in predatory nursing journals,' *Nursing outlook*, 67(6), pp. 664–670. Available at: <https://doi.org/10.1016/j.outlook.2019.05.001>
- Ojala, M. (2025) 'Voices of the searchers: Citations Gone Wild: AI Hallucinations Persist,' *Computers in libraries*, 1 April, p. 31.
- Spennemann, D.H.R. (2025) 'The Origins and Veracity of References 'Cited' by Generative Artificial Intelligence Applications: Implications for the Quality of Responses,' *Publications*, 13(1), p. 12. Available at: <https://doi.org/10.3390/publications13010012>
- Teesside University Student and Library Services (2026) *Beware: Predatory Publishers*. Available at: https://libguides.tees.ac.uk/researchers/predatory_publishers (Accessed: 15 June 2026).
- Ulikozok.com. (2022) 'Two Predatory Publishers from 'Budapest' and 'Great Britain'– Uli Kozok. Available at: <https://ulikozok.com/research/publications/predatory-journals/a-predatory-publisher-from-budapest/> (Accessed: 12 June 2026)
- Walters, W.H. and Wilder, E.I. (2023) 'Fabrication and errors in the bibliographic citations generated by ChatGPT,' *Scientific reports*, 13(1), article 14045. Available at: <https://doi.org/10.1038/s41598-023-41032-5>
- Watson, A.P. (2024) 'Hallucinated Citation Analysis: Delving into Student-Submitted AI-Generated Sources at the University of Mississippi,' *The Serials librarian*, 85(5–6), pp. 172–180. Available at: <https://doi.org/10.1080/0361526X.2024.2433640>
- Wilson, N. (2024) 'Predatory Journals,' *Bioscience*, 74(1), pp. 6–11. Available at: <https://doi.org/10.1093/biosci/biad104>



**Thank you for listening
Any questions?**

Fred Milton f.milton@tees.ac.uk

Teesside Univ. Library colleagues Daisy Campbell, David Watson & Lisa Eveson are thanked for their constructive feedback on this research & presentation.

Copilot (2026) AI-generated image by Copilot with prompt 'produce a landscape cartoon image Roseberry Topping in North Yorkshire', 15 June.